Label Structured Cell Proliferation Models.

H. T. Banks, Frédérique Charles, Marie Doumic-Jauffret, Karyn L Sutton, W Clayton Thompson

Label Structured Cell Proliferation Models

H. T. Banks, Frédérique Charles, Marie Doumic Jauffret,
Karyn L. Sutton and W. Clayton Thompson

Center for Research in Scientific Computation
Center for Quantitative Sciences in Biomedicine
North Carolina State University
Raleigh, NC 27695-8213
and
INRIA Rocquencourt
Projet BANG
Domaine de Voluceau
78153 Rocquencourt, France

June 16, 2010

Abstract

We present a general class of cell population models that can be used to track the proliferation of cells which have been labeled with a fluorescent dye. The mathematical models employ fluorescence intensity as a structure variable to describe the evolution in time of the population density of proliferating cells. While cell division is a major component of changes in cellular fluorescence intensity, models developed here also address overall label degradation.

Key words: Cell proliferation, label structured population dynamics, partial differential equations, moving coordinate systems.

1 Introduction

In [2], a model for tracking labeled proliferating cell populations was effectively used to describe carboxyfluorescein succinimidyl ester (CFSE) labeled lymphocyte cells during proliferation. The model used is given by

$$\begin{aligned} \frac{\partial n(t,y)}{\partial t} &+ \frac{\partial[-cn(t,y)]}{\partial y} = -(\alpha(t,y+ct)+\beta(t,y+ct))n(t,y) \\ &+ \chi_{[y_{\min},y_{\max}-\log\gamma]}(y)2\gamma\alpha(t,y+ct+\log\gamma)n(t,y+\log\gamma). \end{aligned} \tag{1}$$

where y denotes the "true" (Fluorescence-Activated Cell Sorter (FACS) recorded) CFSE fluorescence intensity (FI) (in units of log intensity, log UI) of a cell and $n(t,y)$ is the label-structured population density (cells/UI) of cells with FI y at time t. Here $v(y) = -c < 0$ is the rate of CFSE label loss (due to degradation), $\alpha(t,y+ct)$ and $\beta(t,y+ct)$ are the cell proliferation and death rates, respectively, relative to the moving label coordinate system $y+ct$. Daughter cells arrive at label intensities in the range $y \in [y_{\min}, y_{\max} - \log\gamma]$ and $t > 0$. Because cells naturally lose FI over time [17] even in the absence of division (due at least in part to catabolic activity [10]), the term $v(y) = -c$ represents this natural label loss rate (UI/hr) and hence the FI structure coordinate itself is changing in time. The parameter γ is a label dilution factor, representing the ratio of FI of a mother cell to FI of a daughter cell. Division is coupled with immediate rapid growth of the new daughter cells. The observed or estimated value of γ reflects underlying dynamics (involving mechanisms regulating the growth and division) which occur on a faster time scale and have effectively been integrated over in time in this model.

This model in [2] is an extension of a model first proposed by Bocharov, et al., in [8, 9]. Equation (1) was first heuristically derived in [2] where the emphasis was on using the model with inverse problem methodology to fit experimental data. The purpose of this note is to give a careful and rigorous derivation of a general class of models for proliferation of labeled cells which includes (1) as a special case.

2 Mathematical Model

First, we consider models with total structure label that remains constant in time. Thus in these models there is no label loss. A careful, detailed derivation is given in the Appendix of [2]; the equations are derived there with basic mass balance arguments treating the structure variable as a mass-like quantity. The specific model for the dynamics of life and death processes of a population of cells labeled with CFSE is proposed in [9] as a variation of the Bell-Anderson [3] or Sinko-Streifer [16] (with cell division) population models. The solutions of these models may be directly compared to time snapshots of flow cytometry histograms, which usually depict cells of multiple generations. That is, the models aim to predict the number of cells at a given fluorescence intensity and time, under specified dynamics. Let x denote the CFSE FI (in units of intensity, UI) of a cell and let $n(t,x)$ be the label-structured population density (cells/UI) of cells with FI x at time t. Then for a given label intensity x, the population density is governed by a rate equation involving birth/death and cell division. This can be written as

$$\frac{dn}{dt}(t,x) = -(\alpha(x)+\beta(x))n(t,x) + \chi_{[x_{\min},x_{\max}/\gamma]}(x)2\gamma\alpha(\gamma x)n(t,\gamma x), \tag{2}$$

where $\alpha(x)$ is the cell proliferation rate, $\beta(x)$ is the cell death rate, the daughter cells arrive with intensity x for $x \in [x_{\min}, x_{\max}/\gamma]$ and $t > 0$. The parameter γ is the label dilution factor, representing the ratio of FI of a mother cell to FI of a daughter cell. Division is coupled with immediate rapid growth of the new daughter cells. The observed or estimated value of γ reflects underlying dynamics (involving mechanisms regulating the growth and division) which occur on a faster time scale and have effectively been integrated over in time. As we have noted, a rigorous derivation of this model following mass conservation principles as in the Sinko-Streifer [16] or Bell-Anderson [3] models is presented in the Appendix of [2].

Since the FACS data is typically recorded in terms of *log intensity*, we make the change of variables $z \equiv \log_{10} x$. We use $\tilde{n}$, etc., to denote transformed functions, so $\tilde{n}(t,z) = n(t,10^z) = n(t,x)$. The birth and death functions are simply $\tilde{\alpha}(z) = \alpha(10^z) = \alpha(x)$ and $\tilde{\beta}(z) = \beta(10^z)$, respectively. Their arguments z range over $[z_{\min}, z_{\max}]$ where $z_{\min} = \log_{10}(x_{\min})$, and $z_{\max} = \log_{10}(x_{\max})$. Observe that $\log_{10}(x_{\max}/\gamma) = \log_{10}(x_{\max}) - \log_{10}\gamma = z_{\max} - \log_{10}\gamma$, while $\log_{10}(\gamma x) = \log_{10}\gamma + \log_{10} x = z + \log_{10}\gamma$.

Making the transformation with the new variables as defined, for a fixed z we then obtain the ordinary differential equation model

$$\frac{d\tilde{n}}{dt}(t,z) = -(\tilde{\alpha}(z) + \tilde{\beta}(z))\tilde{n}(t,z) + \chi_{[z_{\min}, z_{\max} - \log_{10}\gamma]}(z) 2\gamma\tilde{\alpha}(z + \log_{10}\gamma)\tilde{n}(t, z + \log_{10}\gamma). \qquad (3)$$

The above equations are adequate in the situation for which total label FI remains constant during the tracking. For many labels, including CFSE, it is known that the label is degraded over time until FI is no longer discernible from background auto fluorescence. (Additional noise in the data is typical for such experiments and is the result of any number of processes, from counting errors to variations in cell shape and size to the functioning of the machine itself [9, 17].) In particular, it is accepted [6, 10, 11, 12] that CFSE FI is lost over time due to catabolic activity within the cell. There may also be label degradation due to photobleaching. Indeed this raises the question of exactly what the reported data represents (label mass, label concentration, etc.) This is difficult to ascertain with any degree of certainty. To better understand and facilitate our subsequent discussions on the parameter γ, we briefly digress to review some salient points on typical data sets and the collection processes used in obtaining them.

As we have already noted above, CFSE histogram data for a population of dividing cells is typically reported in terms of a quantity which is understood to be a "Log of Fluorescence Intensity" (Log FI), measured in unspecified "units of intensity" (UI). However, this data represents an experimental measurement procedure [17] which consists of several complex processes, all of which contribute to an uncertainty with regards to the exact underlying quantity being measured and the process of label degradation. Using hydrodynamic forcing, single cells are pushed through an aperture and into a beam of laser light. Electrons in the CFSE compound absorb some of this light and then emit light of a different wavelength as the electrons jump to an excited state and then fall back to the ground state. This emitted light passes through one or more bandpass filters en route to a detector, which converts the light into a voltage signal. The voltage signal then undergoes an analog-to-digital conversion to produce the data as displayed in the histogram. While the numerous steps of the measurement process are all necessary and are not problematic in themselves, it is unclear at the present how some of these steps directly relate to a conserved quantity (e.g. mass) within the cell population. For instance, it is not immediately known whether the intensity of the absorbed and/or emitted light is proportional to either the mass of CFSE within a cell or to its concentration. Moreover, the mechanisms (cell catabolic activity, photobleaching,

etc.) and their relative importance in label degradation is not completely agreed upon. But we do see significant label degradation in longitudinal data sets observed over sufficiently long periods of time. And data is reported relative to this changing label intensity. The additional steps in the measurement process are sources of uncertainty which can ideally be quantified in the form of a statistical model, but at the moment only further obscure the underlying quantity of measurement and its degradation. We recall that the ratio γ is the ratio $FI_{mother}/FI_{daughter}$ as discussed in [2, 8, 9]. If the measurement is essentially equivalent to one of mass, if all cells divided identically and if the measurement were taken instantaneously upon cell division, then one would expect $\gamma = 2$ (and $\gamma < 2$ would imply creation of label) to provide a best fit to the data. On the other hand, if the measurements are essentially of concentrations, one would expect $\gamma = 1$. The result of [2], where $\gamma = 1.575$ gives a best fit, is a clear indication that the current models average over the cell division and measurement processes. Thus it is likely that more details on the cellular division and measurement processes, as well as some uncertainty across the cell population with respect to these processes, should be introduced in further extensions/refinements of these models if one wishes for a description in terms of more biologically relevant parameters in place of the clearly phenomenological parameter γ.

As explained in [2] and above, it is of interest to introduce a natural label loss velocity v to account for the changing label intensity. That is, the reported intensity coordinate system is actually changing in time (i.e., a moving coordinate system with velocity v). We therefore introduce a change of coordinates involving this velocity to decouple the degradation process from the cell division/proliferation/death processes. Moreover, since the data is recorded relative to this changing structure variable, it is useful to view the proliferation and death rates *relative to this new coordinate system*.

While not common in the biological sciences, it is altogether common in the physical sciences and engineering to consider velocities (i.e., rates of change) relative to different coordinate or reference frames. For example, in the mechanics and motion of continua (elasticity and fluids) and deformable bodies [1, 4, 5, 13, 14], it is frequent to encounter velocities relative to a *fixed* coordinate system (in a *Lagrangian* formulation) or relative to a *moving* coordinate system (in an *Eulerian* formulation). One description for motion is made in terms of the material or fixed referential coordinates, and is called a material description or the *Lagrangian description*. In this formulation, an observer standing in the fixed referential frame observes the changes in the position and physical properties as the material body moves in space as time progresses. This formulation focuses on individual particles as they move through space and time. The other description for motion is made in terms of the spatial or current coordinates, called a spatial description or *Eulerian description*. The coordinate system is relative to a moving point in the body and hence is a *moving coordinate system*. An intuitive comparison of these two descriptions would be that in the Eulerian description one places the coordinate or reference system for motion of an object *on* the object as it moves through a moving fluid (e.g., on a boat in a river) while in the Lagrangian description one observes and describes the motion of the object from a fixed vantage point (e.g., motion of the boat from a fixed point on a bridge over the river or on the side of the river.).

Motivated by the above discussions, we introduce a formal (implicit) change of variables $y = z + vt$ where $v < 0$ is label degradation velocity. Because we really don't understand completely the degradation process (there appears to be little agreement as to what variables on which this velocity might depend) and to allow for generality (other labels that might be used may well degrade in different ways due to different mechanisms), we allow this velocity to depend on time as well as the "true" current label intensity y. Thus the change of variables is actually $y = z + v(t, y)t$. We tacitly

assume that this can be inverted to solve for $y = \mu(t,z)$ explicitly in $\mathcal{F}(t,z,y) = y - v(t,y)t - z = 0$. Global Implicit Function Theorems [15] provide conditions (including $\frac{\partial \mathcal{F}}{\partial y} = 1 - \frac{\partial v}{\partial y} t \neq 0$) for this to hold. Observe that this condition holds trivially for v constant (as in [2]) or only time dependent. Using this change of variables in (3), we obtain a generalization of equation (1). To see this, under the necessary assumptions, we can solve $y - v(t,y)t - z = 0$ for $y = \mu(t,z)$ so that $\mu(t,z) - v(t,\mu(t,z))t - z = 0$ and hence

$$\frac{\partial \mu}{\partial t}(t,z) = \left[1 - \frac{\partial v}{\partial y}(t,y)t\right]^{-1} \left[v(t,y) + \frac{\partial v}{\partial t}(t,y)t\right].$$

Defining $\hat{n}(t,y) = \hat{n}(t,\mu(t,z)) = \tilde{n}(t,z)$ and observing that

$$\frac{d\tilde{n}(t,z)}{dt} = \frac{\partial \hat{n}}{\partial t}(t,y) + \frac{\partial \hat{n}}{\partial y}(t,y)\frac{\partial \mu}{\partial t}(t,z),$$

we may use equation (3) to find

$$\begin{aligned} \frac{\partial \hat{n}(t,y)}{\partial t} &+ \left[1 - \frac{\partial v}{\partial y}(t,y)t\right]^{-1} \left[v(t,y) + \frac{\partial v}{\partial t}(t,y)t\right] \frac{\partial \hat{n}(t,y)}{\partial y} \\ &= -(\tilde{\alpha}(y - v(t,y)t) + \tilde{\beta}(y - v(t,y)t))\hat{n}(t,y) \\ &+ \chi_{[y_{\min}(t,\gamma), y_{\max}(t,\gamma)]}(y) 2\gamma\tilde{\alpha}(t, y - v(t,y)t + \log\gamma)\hat{n}(t, y + \log\gamma), \end{aligned} \tag{4}$$

where

$$y_{\min}(t,\gamma) \equiv \min_{z \in [z_{\min}, z_{\max} - \log\gamma]} y(t,z), \quad y_{\max}(t,\gamma) \equiv \max_{z \in [z_{\min}, z_{\max} - \log\gamma]} y(t,z).$$

Thus the range $[z_{\min}, z_{\max} - \log\gamma]$ for z transforms into a range of $[y_{\min}(t,\gamma), y_{\max}(t,\gamma)]$ for y and hence the χ term in (3) transforms into the one of (4).

We observe that this "Eulerian" formulation depends explicitly on the velocity v of the coordinate system. This can be compared to equations (2.21) (the Eulerian formulation) and equations (2.24) (the Lagrangian formulation) in [1] where the former also depends explicitly on the velocity of the coordinate system.

We observe that in the case the velocity is constant (i.e., $v = -c$ as in [2]), then equation (4) reduces to precisely the equation (14) of [2] or equation (1) above in the case that α, β are also assumed to depend explicitly on time t.

Acknowledgements

This research was supported in part by Grant Number R01A I071915-07 from the National Institute of Allergy and Infectious Diseases and in part by the Air Force Office of Scientific Research under grant number FA9550-09-1-0226.

References

[1] H. T. Banks, Shuhua Hu and Zackary R. Kenz, A brief review of elasticity and viscoelasticity, CRSC-TR10-08, N.C. State University, May, 2010; *Advances in Applied Mathematics and Mechanics,* submitted.

[2] H. T. Banks, W. Clayton Thompson, Karyn L. Sutton, Gennady Bocharov, Dirk Roose, Tim Schenkel, and Andreas Meyerhans, Estimation of cell proliferation dynamics using CFSE data, CRSC-TR09-17, August, 2009; *Bull. Math. Biology*, (2010), to appear.

[3] G. Bell and E. Anderson, Cell growth and division I. A mathematical model with applications to cell volume distributions in mammalian suspension cultures, *Biophysical Journal,* **7** (1967), 329–351.

[4] Y. C. Fung, *A First Course in Continuum Mechanics*, Prentice Hall, Englewood Cliffs, NJ, 1994.

[5] Y. C. Fung, *Biomechanics: Mechanical Properties of Living Tissue*, Springer-Verlag, Berlin, 1993.

[6] E. D. Hawkins, Mirja Hommel, M. L Turner, Francis Battye, J. Markham and P. D Hodgkin, Measuring lymphocyte proliferation, survival and differentiation using CFSE time-series data, *Nature Protocols,* **2** (2007), 2057–2067.

[7] T. Luzyanina, S. Mrusek, J. T. Edwards, D. Roose, S. Ehl and G. Bocharov, Computational analysis of CFSE proliferation assay, *J. Math. Biol.,* **54** (2007), 57–89.

[8] T. Luzyanina, D. Roose and G. Bocharov, Distributed parameter identification for a label-structured cell population dynamics model using CFSE histogram time-series data, *J. Math. Biol.,* to appear.

[9] T. Luzyanina, D. Roose, T. Schenkel, M. Sester, S. Ehl, A. Meyerhans and G. Bocharov, Numerical modelling of label-structured cell population growth using CFSE distribution data, *Theoretical Biology and Medical Modelling,* **4** (2007), 1–26.

[10] A. B. Lyons, Divided we stand: tracking cell proliferation with carboxyfluorescein diacetate succinimidyl ester, *Immunology and Cell Biology,* **77** (1999), 509–515.

[11] A. B. Lyons and K. V. Doherty, Flow cytometric analysis of cell division by dye dilution, *Current Protocols in Cytometry*, (2004), 9.11.1-9.11.10.

[12] A. B. Lyons and C. R. Parish, Determination of lymphocyte division by flow cytometry, *J. Immunol. Methods,* **171** (1994), 131–137.

[13] J. E. Marsden and T. J. R. Hughes, *Mathematical Foundations of Elasticity*, Dover Publications, Inc., Mineola, NY, 1994.

[14] R.W. Ogden, *Non-Linear Elastic Deformations*, Dover Publications, Inc., Mineola, NY, 1984.

[15] I. W. Sandberg, Global implicit function theorems, *IEEE Trans Circuits and Systems*, **CAS-28** (1981), 145–149.

[16] J. Sinko and W. Streifer, A new model for age-size structure of a population, *Ecology,* **48** (1967), 910–918.

[17] Wikipedia, http://en.wikipedia.org/wiki/Fluorescence_spectroscopy.